



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE United States Patent and Trademark Office Address: COMMISSIONER FOR PATENTS P.O. Box 1450 Alexandria, Virginia 22313-1450 www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/157,884	09/21/1998	ANDRES VEGA-GARCIA	777.179US1	3059
23460	7590 05/07/2004		EXAMINER	
LEYDIG VOIT & MAYER, LTD TWO PRUDENTIAL PLAZA, SUITE 4900 180 NORTH STETSON AVENUE CHICAGO, IL 60601-6780			PRIETO, BEATRIZ	
			ART UNIT	PAPER NUMBER
			2142	οi
			DATE MAILED: 05/07/2004	3

Please find below and/or attached an Office communication concerning this application or proceeding.

U.S. Patent and Trademark Office PTOL-326 (Rev. 1-04)

1) Notice of References Cited (PTO-892)

Paper No(s)/Mail Date

2) Motice of Draftsperson's Patent Drawing Review (PTO-948)

3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)

4) Interview Summary (PTO-413)

Paper No(s)/Mail Date.

6) Other:

5) Notice of Informal Patent Application (PTO-152)

DETAILED ACTION

- 1. This communication is in response to Request for reconsideration filed 2/24/04, claims 1-2, 4-8, 18-26, 28-29 and 31-42 remain pending in this application.
- 2. Objection with respect to claim 22 and 23 is withdrawn, both claims are directed to the same statute for implementing the same apparatus claim 21, the former claim comprising the instruction for implementing the apparatus and the later claim for implementing the process limitations
- 3. Quotation of 35 U.S.C. §103(a) which forms the basis for all obviousness rejections set forth in this Office action may be found in previous action.
- 4. Claims 1, 18, 21, and 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Smith et. al. U.S. Patent No. 6,128,649 (Smith hereafter) in view of ITU-T H.323 Centralized multipoint configuration/clarification, Northlich, B., Onlive Technologies, Inc., Feb. 1997 (Northlich hereafter).

Regarding claim 1, Smith substantial features of the invention as claimed, teaching a system/method in a network conferencing environment (Smith: abstract, col 1/lines 64-67) for delivering a plurality of video or audio media data type signals (Smith: Figs. 1-5a, audio and/or video media streams, col 3/lines 20-35) the system comprising;

transmitting a set of media data streams on to the network, set of media data streams generated from the plurality of video or audio type signals (Smith: abstract, video streams, distributing audio/video streams across the network, col 1/lines 53-67);

transmitting means include means for removing silences from said data streams of the audio signals transmitted by the transmitter (Smith: identifying silence stream, col 9/lines 5-9, removing said identified streams from data audio transmission stream by closing audio channel from originator);

a receiver for receiving the set of data stream from the network (Smith: col 1/lines 30-33, 53-57, 63-co 2/line 2, 25-27, col 4/lines 2/lines 30-36, 40-55);

the receiver including a selectively routing, filtering or separating media streams (i.e. demultiplexing) means (Smith: 1 of Fig. 21 for multiplexing, col 1/lines 53-62) for dynamically selecting a subset of the set of data streams (dynamic selection (13 of Fig. 21)) (Smith: col 6/lines 49-col 7/line 26, dynamic selection of multiple media streams, see abstract, and multiplexing means col 27/lines 31-55);

two or more receiver media data stream (payload) handler modules (Smith: col 20/line 11-col 21/line 25, reception processing modules, i.e. receiver, reception audio/video process modules, i.e. two, col 7/lines 35-48);

two receiving (Smith: media-in portion 20 of Fig. 5a, col 7/lines 35-39) including receivers coupled to said demultiplexing means for handling routed data streams (Smith: first reception means col 7/lines 58-67, having decoding (28) means, and second reception means col 8/lines 12-22);

two decoder modules coupled to the demultiplexing means for decoding routed data streams (Smith: col 20/lines 11-30) two or more type of data streams (Smith: first decoder see col 7/lines 58-67, second decoder see col 22/lines 61-67 associated with respective media type data processing modules (26/32 of Fig. 5b); although the above-mentioned prior art teach dynamically selection a subset of the subset of data stream, Smith does not explicitly teach wherein the selection is based on a source identifier and a payload type;

Northlich discloses a clarification to ITU-T recommendation H.323, this recommendation describes multiple terminals supporting the transmission by transmitters of multimedia type of data streams on the network and the reception by receivers of selectively parceled multimedia type of data streams in a multipoint conferencing environment, wherein all terminals support different media types those recommended in the H.323. Northlich clarifies a switch process pertaining the handling of data streams, disclosing demultiplexing data streams (RTP) based on SSRC and payload type, to include streams of video and/or audio channels in a conference environment (see page 2).

It would have been obvious to one ordinary skilled in the relevant art at the time the invention was made giving the suggestion of Smith for combining multiple audio/video streams received from multiple participant's in an conferencing network. One would have look at prior pertaining the delivery, separation, processing and rending of combined multimedia stream to multiple conferee recipients in a conference network. Northlich discussing pertaining standard technology in conferencing environment includes means for selecting a subset of the set of data stream based on a source identifier and a payload, as taught by Northlich. Combined teachings would enable one ordinary skilled to separate and route, i.e. multiplex set of data streams received based on source identifier and payload type to corresponding subsequent post reception processes such as forwarding to corresponding pre-rendering processors, e.g. corresponding codecs, motivation would be enable a terminal having different audio/video capabilities to support simultaneous session in multiple data stream types to mix both audio and/or video in a conferencing system.

Regarding claim 18, further teach a method of conducting a network conference with two or more computer systems (Smith: users 3 of Figs. 1-3 illustrating a conferencing network), comprising:

modules (Smith: modules 33-35 of Fig. 5B) for monitoring incoming audio data for each of a plurality of conference parties for active or inactive status (Smith: determining active media streams col 3/lines 61-col 4/line 11, with stream activity monitoring/detection means (33), determine state active change, col 9/lines 10-57, determining which said streams are silent or less active, Figs. 7-8);

monitoring incoming audio or video for a new speaker (Smith: new stream activity detection means, col 10/lines 44-col 1/line 19, stream activity associated with conference participant's GUI event (60, 70), col 9/lines 10-57);

replacing audio data having the inactive status with data of the new speaker (Smith: means for substituting a set of data from another (third) conference participant with data set from a respective determined inactive participant, comprising means for determining (150) the most silent stream to be replace, wherein in response to a positive determination replacing (dropped) said most silent stream with said (third) stream, col 10/line 43-col 11/line 19, replacing a silent stream with a another (third) data set associated with another participant, mean for detecting most recent speaker and performing substitution steps, col 19/lines 3-45);

receiving audio or video from first and second computer system (Smith: reception modules 28/34 for receiving payload streams from network see col 20/lines 11-23 from users on respective systems see col 1/lines 63-col 2/lines 6, Figs. 1-3 users 3);

routing the audio or video data to respective decoder based on determined audio or video payload type of the audio or video data stream and a source identifiers (Northlich; page 2).

Regarding claim 21, prior art teaches in a conferencing system:

receiving the set of data stream from the network operating under RTP, i.e. "RTP compliant data stream" (Smith: stream reception from the network see col 20/lines 11-23, reception/transmission processes are RTP compliant see col 21/lines 22-50);

dynamically selecting a subset of the set of data streams (Smith: dynamic selection means, col 6/lines 49-col 7/line 26, dynamic selection of multiple media streams, see abstract, and multiplexing means col 27/lines 31-55);

routing RTP data stream(s) based on payload type(s) and a source identifier (Northlich: page 2);

two or more receiver media data stream (payload) handler modules (Smith: col 20/line 11-col 21/line 25, reception processing modules, reception audio/video process modules, col 7/lines 35-48); specifically in regards to claim 21;

two decoder modules coupled to the demultiplexing means for decoding routed data streams (Smith: video decoder see col 7/lines 58-67 and 23/lines 1-11, media type based decoding see col 20/lines 11-30, audio decoder see col 22/lines 61-67);

a rendering means coupled to the decoder for playing back one RTP data stream (Smith: presentation of the resulting stream (after decoding stream) to output device for displaying, i.e. rendering or "playback" see col 20/lines 39-42, selecting streams from users 1,2, & 3 for display, see col 2/lines 25-27, for display on two or more user terminals see col 4/lines 2/lines 30-36, 40-55).

Regarding claim 24, the combined teachings as discussed above, teaches a network conferencing system comprising:

receiving means (Smith: 26/32 of Fig. 5b) for receiving via a communication network respective first and second sets of data of at least one payload type from respective first and second conference participant (Smith: col 1/lines 53-57, 63-col 2/line 2, 25-27);

first/second decoder for decoding payload type(s) of data (Smith: col 7/lines 58-67, col 22/lines 61-col 23/line 11 audio/video decoders);

means (e.g. demultiplexer) for routing data said received data to first or second decoder (Smith: media-in portion 20, col 7/lines 35-39, 58-67, col 8/lines 12-22);

two decoder modules coupled to the demultiplexing means for decoding routed data streams (Smith: col 20/lines 11-30) two or more type of data streams (Smith: first decoder see col 7/lines 58-67, second decoder see col 22/lines 61-67 associated with respective media type data processing modules (Smith: modules 26/32 of Fig. 5b) based on payload type and at least one source identifier (Northlich; page 2);

means (Smith: stream activity monitoring/detection modules 33-35 of Fig. 5b) include determining whether a set of data is associated with an inactive conference participant (Smith: determine if stream i.e. "sets of data" associated with a conference participant is active see col 3/lines 61-col 4/line 11, determine activity see col 9/lines 10-57);

means responsive to determination of the inactive conference participant, for substituting a third set of data from a third conference participant, for at least the one of the first and second sets of data associated with the inactive conference participant (Smith: substitution see col 10/line 43-col 11/line 19, replacing a silent stream with another speaker and performing substitution steps, col 19/lines 3-45).

5. Claims 2, 4-8, 19-20, 22-23, 25-26, 28-29, and 31-42 are rejected under 35 U.S.C. 103(a) as being unpatentable over Smith in view of Northlich in further view of H.323 ITU-T: Audiovisual and multimedia systems, Nov. 1996, pages 1-71 (referred to as H.323 hereafter).

Regarding claims 2, 4 and 8, combined teachings as discussed above, the neither Smith nor Northlich reference do not explicitly teach wherein the audio decoders are particularly of G.711 and G.723.1.

H.323 this recommendation describes multiple terminals supporting the transmission by transmitters of multimedia type of data streams on the network and the reception by receivers of selectively parceled multimedia type of data streams in a multipoint conferencing environment, wherein all terminals support different media types those recommended in the H.323 series (see summary on page (i), section 6.2-6.2.2 on page 11, Audio codec sec 6.2.5 on page 13). One channel for each type of media data stream type includes audio codes, G.7.11, G.722, G.728 and G. 723 (see page (i), Fig. 4 of section 6.2).

It would have been obvious to one ordinary skilled in the art at the time the invention was a made given the suggestion of Smith for using codec for encoding/decoding audio and video associated with respective audio/video reception processing modules (i.e. "type based payload handles") and the clarification of Northlich. One ordinary skilled in the art would have look at pertinent art directed to the processing of audio/video data in a conferencing environment. Audio Codecs particularly of G.711 and G.723, are known as standard. One ordinary skilled will be motivated to utilize audio decoders are particularly of G.711 and G.723 enhance the capabilities of the conference system enabling participants of different capabilities to communicate.

Regarding claim 5, the combined teachings as discussed above however do not explicitly teach for mixing an audio stream operatively coupled to the two or more corresponding decoders.

Official Notice (see MPEP § 2144.03 Reliance on "Well Known" Prior Art) is taken that a mixer was old and well known in the Data Processing art. It would have been obvious to one of ordinary skill in the art at the time of applicant's invention to include a mixer for mixing audio stream, motivation would be to render a composite audio signal to the user.

Regarding claim 6, media rendering module operatively coupled to the decoder(s) (Smith: presentation of the resulting stream (after decoding stream) on an output device see 20/lines 39-42).

Regarding claim 7, wherein data processor(s) "payload handler(s)" includes: means for combining two or more data packets (Smith: mixing see col 1/lines 30-35, mixing see col 1/lines 63-col 2/line 2).

Regarding claim 19, decoding the audio or video data from the first and second computer systems (Smith: data received from two or more user terminals see col 1/lines 53-57, 63-col 2/line 2, decoding audio or video, video decoders see col 7/lines 58-67, media type decoding see col 20/lines 11-30, audio decoder see col 22/lines 61-67, video decoder see col 23/lines 1-11);

rendering the audio or video data from the first and second computer systems (Smith: selecting streams from users 1,2, & 3 for display, i.e. rendering see col 2/lines 25-27, for display on two or more user terminals see col 4/lines 2/lines 30-36, 40-55).

Regarding claim 20, the claim is substantially the same as claim 2, same rationale is applicable.

Regarding claims 22-23, a machine-readable medium comprising instruction for implementing the modules (Smith: software implementation of disclosed method see col 4/lines 37-39).

Regarding claim 25, this method claim comprises the combination of limitations claims 1, 4, 18-19, 21 and 24, as discussed above, same rationale of rejection is applicable.

Regarding claim 26, means (32 of Fig. 5b) receiving a plurality of audio data streams from a corresponding plurality of conference participants (Smith: col 1/lines 30-33, col 1/line 53-col 2/line 2);

means for selecting a subset of plurality of audio data streams of different types (Smith: selecting subset of data streams col 6/lines 49-col 7/line 26, dynamic selection of multiple media streams, see abstract, and multiplexing means col 27/lines 31-55;

audio payload of different types associated with respective encoding type (H.323: see summary on page (i), section 6.2-6.2.2 on page 11, Audio codec sec 6.2.5 on page 13), one channel for each type of media data stream type includes audio codes, G.7.11, G.722, G.728 and G. 723 see page (i), Fig. 4 of section 6.2);

means for routing data received by said receiving means to the first or the second decoder module based on the payload type and at least one source identifier (Northlich: page 2);

means for rendering the selected subset of audio data streams (Smith: audio streams are sent to the users see col 1/lines 63-col 2/line 2, a single audio output is provided to the user from all input audio streams see col 8/lines 12-22, rendered audio see col 18/lines 17-23).

Regarding claim 28, wherein the selected subset of audio data stream includes a first audio data stream and a second audio data stream (Smith: selection means 13 of Fig. 21, col 6/lines 49-col 7/line 26, see abstract, and multiplexing means col 27/lines 31-55), and wherein the system further comprises:

means (Smith: 33 of Fig. 6) for determining whether one or more of the first and second audio data streams is associated with an inactive conference participant (Smith: determining activity of streams see col 3/lines 61-col 4/line 11, using stream activity monitoring/detection means (33), see col 9/lines 10-57, determining streams are silent or less active, Figs. 7-8);

means, responsive to determination of the inactive conference participant, for substituting a third audio data stream from a third conference participant, for at least the one of the first and second audio data streams associated with the inactive conference participant (Smith: means for substituting a set of data from another (third) conference participant with data set from a respective determined inactive participant, comprising means for determining (150) the most silent stream to be replace, wherein in response to a positive determination replacing (dropped) said most silent stream with said (third) stream, col 10/line 43-col 11/line 19, replacing a silent stream with a another (third) data set associated with another participant, mean for detecting most recent speaker and performing substitution steps, col 19/lines 3-45).

Regarding claim 29, this claim comprises limitations that are substantially the same as claim 26, same rationale is applicable.

Regarding claim 31, this claim is comprises limitations that are substantially the same as claim 28, same rationale of rejection is applicable.

Regarding claim 32, this claim comprises the combined limitation of claims 26, and 28-29 same rationale of rejection is applicable.

Regarding claim 33, this claim comprises limitations that are substantially the same as combined claim 27, same rationale is applicable.

Regarding claims 34-36, the combined teachings as discussed above, further teach

wherein the selected subset includes a first video data stream formatted according to a first protocol an a second video data stream formatted according to a second protocol (H.323: summary page

Application/Control Number: 09/157,884 (VEGA-GARCIA et. al.) Art Unit: 2142

(i) and sections 6.2-6.2.2 on page 11), wherein the data streams in the selected subset are most recently activate data streams (Smith: col 1/lines 63-col 2/line 2),

selection based on monitored activity, event detection data stream activity (Smith: col 19/lines 3-21, most recent audio data stream activity associated with a participant, col 19/lines 22-45, wherein the first and second sets of data streams are audio signal data of a multicast group of (e.g. dialogue) between two or more participants).

Regarding claims 37-42, synchronization source identifier (Northlich: col 13/lines 3-24).

Response to arguments

6. Regarding claims 22 and 23, applicant indicated that these claim are of different scope.

In response to the above assertion, it is respectfully noted that both claims of the same statute are implementations of the same apparatus claim 21, the former claim comprising the instruction for implementing the apparatus, and the later claim for implementing the process limitations thereof, there are not considered to be of different scope.

Regarding claims 1, 18, 21 and 24, applicant argues, prior art does not teach removing silence or background from the data streams of audio signals transmitted by said at least one transmitter, because according to applicant, claim limitation indicates the removal from the transmitted data stream silences or background forming part of the data stream while the data stream sans silences or background still continue to be transmitted by the transmitter, meaning that the data stream is not switched or shut-off, because the Smith reference according to applicant's interpretation, removes the identified streams from the audio transmission by closing that audio channel, thereby as applicant further interprets, the prior art removes the entire data stream from one of the conference participants from the audio channel if the conference participant is currently silent.

In response to the above-argument, portions cited by applicant have been reviewed, however none disclose removing the entire data stream, nor shutting off a data stream transmitted. In this case, col 9/lines 5-9, discusses the five threads of the MDM cover the response to GUI events, a periodic consistency check to see if additional streams can be displayed, and responses to new T_ stream activity or silence, and the response to a closing of a D-stream by its originator, where D-stream is a Display stream see col 7/lines 58-67 and MDM is decision module for implementing dynamic selection control functions see col 8/lines 23-28, further col 9/lines 34-36 discuss where "if a violation of constraints is detected, the thread tries to remove display streams corresponding to silent audio channel, until

Application/Control Number: 09/157,884 (VEGA-GARCIA et. al.)

Art Unit: 2142

conformance with the constraint is achieved" and furthermore col 10/lines 19-22, 25-27 discuss and "removing from display the D_stream with the most silence in its corresponding audio stream among the showing congestion", and "removing from display the D_stream with the most silence in its corresponding audio stream. The portions noted by applicant provide not evidence of applicant's conclusion, specifically, do not disclose where the prior art removes the entire data stream from one of the conference participants from the audio channel if the conference participant is currently silent.

It is further noted that the features upon which applicant relies (i.e., removing from the transmitted data stream silences or background forming part of the data stream while the data stream sans silences or background still continues to be transmitted by the transmitter, meaning that the data stream is not switched or shut-off") are not recited in the rejected claim(s). This is not a suggestion of any sort. Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993). Arguments are not persuasive.

8. Applicant argues prior does not teach claim limitation as recited, specifically, a demultiplexer for dynamically selecting a subset of the set of data streams, because the portion cited by examiner, according to applicant describes a multicast control unit that multiplexes or mixes and distributed video streams transmitted by the network to/from users forming a centralized topology. This according to applicant describes any demultiplexing of the mixed signal at any one of the user to which the mixed data stream is transmitted. Applicant indicated that cited portion indicated that separate audio and video streams are sent to each user, thereby the is no need of a demultiplexer according to applicant's interpretation of the prior art. Further, the dynamic selection controller according to applicant does not perform the functions of a demultiplexer because it does not receive any data streams whatsoever.

In response to the above-mentioned argument, claim limitation recites, "a demultiplexer for dynamically selecting a subset of the set of data streams (i.e. video or audio signals)". Applicant assertion that the MCU does not receive any data streams is noted. However, Fig. 1 illustrates an MCU (1) module receiving media streams from users and outputting a mixed/selected stream, description of said figure discloses that the "MCU multiplexes or mixes and distributes video stream and... is an expensive dedicated selecting equipment" (see col 1/lines 53-61). Further disclosing in regards to said figure, that "each user sends its own video and audio to the MCU, when controlled by a chairman, the MCU selects one of the incoming video streams... the input audio streams with the most activity could be selected for mixing and outputting" (see col 1/lines 63-col 2/line 3). This centralized switching approach involves selecting one of the streams from users 1, 2, 3 for display (see col 2/lines 24-28).

Figure 2 illustrates a select video stream transmitted from the network (11) received by users terminal (10), which streams to be selected is determined *dynamically* and *only the selected streams* are passed to the user. Dynamic *selection from multiple streams* enables the user to concentrate on the content (see abstract).

Arguments that prior art does not teach dynamically selecting a subset of the set of data streams are not persuasive.

9. THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a). A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Prosecution of this application is closed by means of this final office action § 1.113, applicant may request continued examination of the application by filing a Request for Continued Examination of under 37 CFR § 1.114 and providing the corresponding fee set forth in § 1.17(e) for the submission of, but not limited to, new arguments, an information disclosure statement, an amendment to the written description, claims, drawings, or new evidence in support of patentability. Or applicant whose claims has been twice rejected, may appeal from the decision of the administrative patent judge to the Board of Patent Appeals and Interferences under 35 U.S.C. §134.

Application/Control Number: 09/157,884 (VEGA-GARCIA et. al.)

Art Unit: 2142

Page 11

SUPERVISORY PATENT EXAMINER

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Prieto, B. whose telephone number is (703) 305-0750. The Examiner can normally be reached on Monday-Friday from 6:00 to 3:30 p.m. If attempts to reach the examiner by telephone are unsuccessful, the Examiner's Supervisor, Jack B. Harvey can be reached on (703) 305-9705. Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-3800/4700.

Any response to this final action should be mailed to:

Box AF

Commissioner of Patents and Trademarks Washington, D.C. 20231

or faxed to the Central Fax Office: (703) 872-9306, for Official communications and entry

Or Telephone:

(703) 306-5631 for TC 2100 Customer Service Office

Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Drive, Arlington VA, Sixth Floor (Receptionist).

B. Prieto

TC 2100 Patent Examiner

May 5, 2004